P.c.b. layout for high-speed Schottky t.t.l.

Requirements of printed-board design for low inductance and effective decoupling

by D. Walton, B.Sc. (Hons), Ph.D.

A great deal has been written on the subject of logic design and quite comprehensive books appear almost monthly. In general, however, the published material neglects an extremely important area and one which probably gives the most trouble to practising engineers. This area, which is dealt with in the present article, is concerned with the layout of logic on printed circuit boards in order to ensure reliable operation. The impetus for writing this article comes from the author's own experience of the lamentable lack of understanding of these basic considerations.

IT SHOULD not be concluded from the preamble that the subject is a difficult one; indeed the mathematics employed in the present paper is extremely elementary. The problems are caused rather by the historical progression from analogue to digital techniques with the consequent carrying out of well-tried analogue practices into the digital environment. Unfortunately, the requirements for digital circuitry are frequently opposite to those needed by the analogue variety and hence there is a need for a complete reconsideration of the requirements.

Low inductance bussing

To understand the criteria which determine how the supply and GND lines should be distributed to the t.t.l., first take the case of a t.t.l. gate driving its output line from low to high. For the gate to drive the output line high it must pass current into it. The output line must be considered as a transmission line of impedance Z_0 if its length exceeds 10cm. In practice, Z_0 will be in the region of 100Ω and for a single logic signal changing from low to high the instantaneous output current will be given by $I_0 = 5/100 = 50$mA. This current must be obtained from the supply rails in a time comparable to the risetime of the signal. If, for Schottky t.t.l., $t_{r(min)} \simeq 1.5$ns, then charge must be transferred from the decoupling capacitor to the gate and hence to the output line in this time. Remember that charge is obstructed from flowing into the gate by the inductance, L, of the loop ABCD in Fig. 1. If this is approximately 2cm square with reasonable track width then, using the formula for parallel wires, $L = \ln(a/r)\,\mu_0/4\pi \approx 30$nH. The e.m.f. dropped across L will then be given by $E = -L di/dt$. Therefore,

$$E = \frac{30\times10^{-9}\times50\times10^{-3}}{1.5\times10^{-9}}$$

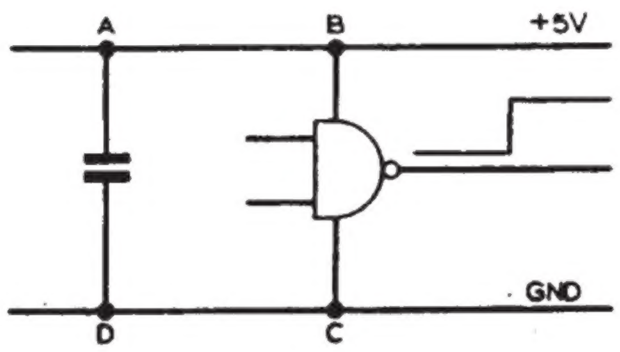


Fig 1. Example of gate, with decoupling, producing a low-to-high transition.

$= 1$ volt

This is a considerable voltage and it should be remembered that it is the result of a single gate switching. If all four gates in a pack switch together the currents will be additive and the rail will fall by 4 volts.

The first requirement of a power distribution system must therefore be low inductance between the i.c. and the decoupling capacitor. This is achieved by the track layout shown in Fig. 2(b), where a low inductance path from C to the i.c. is provided by keeping the V_{CC} and GND tracks close together.

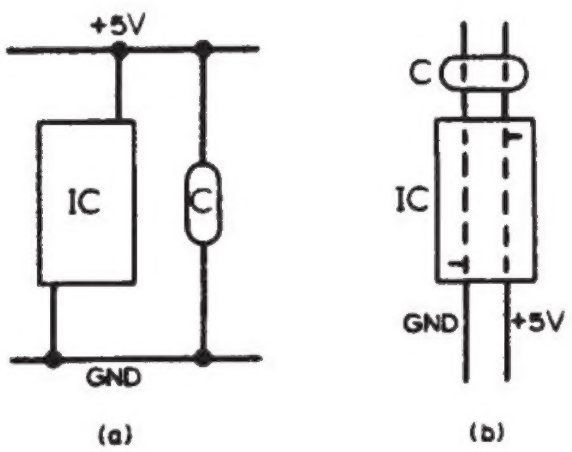


Fig. 2. Two ways of laying out supply lines. Preferred method, giving lower inductance, is at (b).

Manufacturers of i.cs usually specify one decoupling capacitor for every 5–10 i.cs which, with the track layout of Fig 2(a) results in prohibitively high inductance between the capacitor and the worst-case positioned i.c. The safest course is to provide the track layout as in Fig. 2(b) but also to put one capacitor adjacent to each i.c. Clearly, this can be achieved by having one capacitor for each pair of i.cs.

Decoupling capacitors

The foregoing argument shows that the capacitor is better thought of as a reservoir capacitor which supplies the local, instantaneous current demands as i.cs switch. This means that the important parameter for such a capacitor is the instantaneous current which it can supply. Some manufacturers specify capacitors for i.c. decoupling by giving the maximum pulse risetime, which corresponds to a maximum current for a given size of capacitor. For instance, a 47nF capactior specified at 50V/μs can supply a current given by

$$i = C.\frac{dv}{dt} = 47 \times 10^{-9} \times \frac{50}{10^{-6}}$$
$$= 2.5\text{A},$$

which is adequate in the context of the previous calculation.

The other check to make is that the current drawn from the capacitor does not cause its voltage and hence the rail voltage to fall excessively. If the local demand is equal to 10 gates switching, the current demand will be 500mA; to be safe, assume that this demand lasts for 10ns, and design for a voltage drop at the capacitor of 50mV.

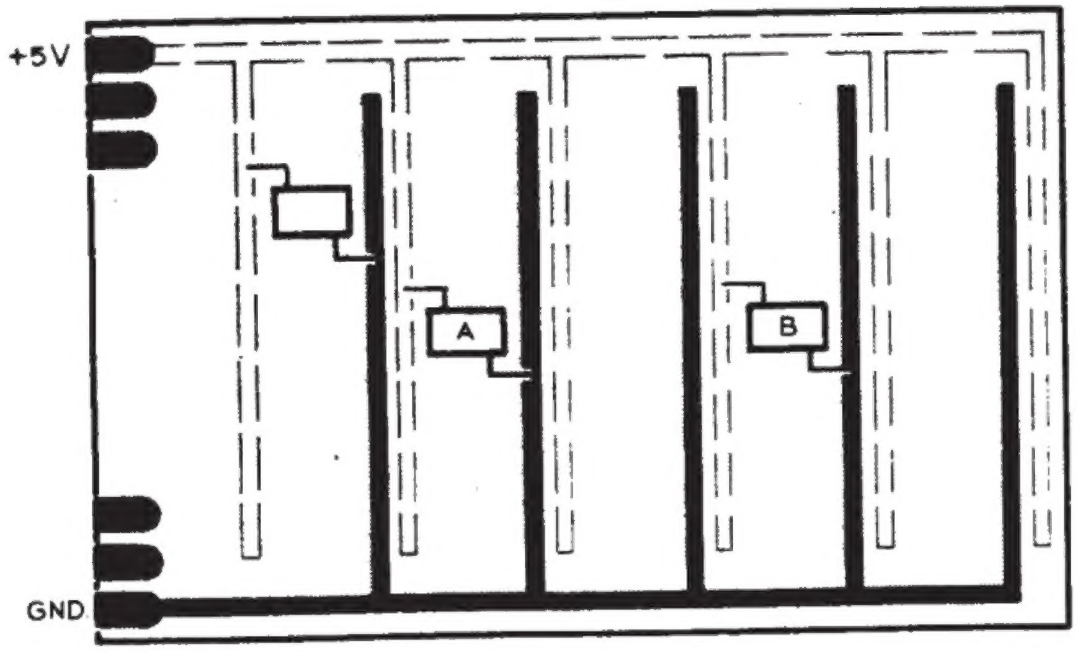


Fig. 6. A bad layout giving high inductance and few adjacent signal return paths, which leads to cross-talk.

sequent reduction in the speed of the system. As was explained earlier, the layout of Fig 6 is also bad from the point of view of placing excessive inductance in the way of charge travelling between i.cs and decoupling capacitors.

Recommended layout

A recommended scheme for laying out a printed circuit board is shown in Fig. 7. The power rails are run as close together as possible along the columns of integrated circuit packages and are interconnected at the top and bottom of the board. These provide return paths for logic signals travelling parallel to them. To provide return paths for signals travelling across the board the ground pins of the packages are connected together from left to right. Thin track, of the same thickness used for signal interconnexions can be used for this. A tantulum bead 10μF decoupling capacitor is provided between each pair of i.cs. Notice also that ground connexions are brought out at regular intervals across the edge connector. These provide return paths for signals travelling on and off the board.

If all these design rules are followed a reliable system will result and the consequent savings in servicing and testing will amply repay a little consideration given to board layout at the design stage.

References

1. Bonham, B. 'Schottky t.t.l.', T.I. application report B93.
2. Catt, I. 'Crosstalk (noise) in digital systems', I.E.E.E. Transactions on Electric Components', EC-16, 743-763 (1967).
3. Gunston, M. A. R. 'Microwave transmission live impedance data'. Van Nostrand Reinhold.

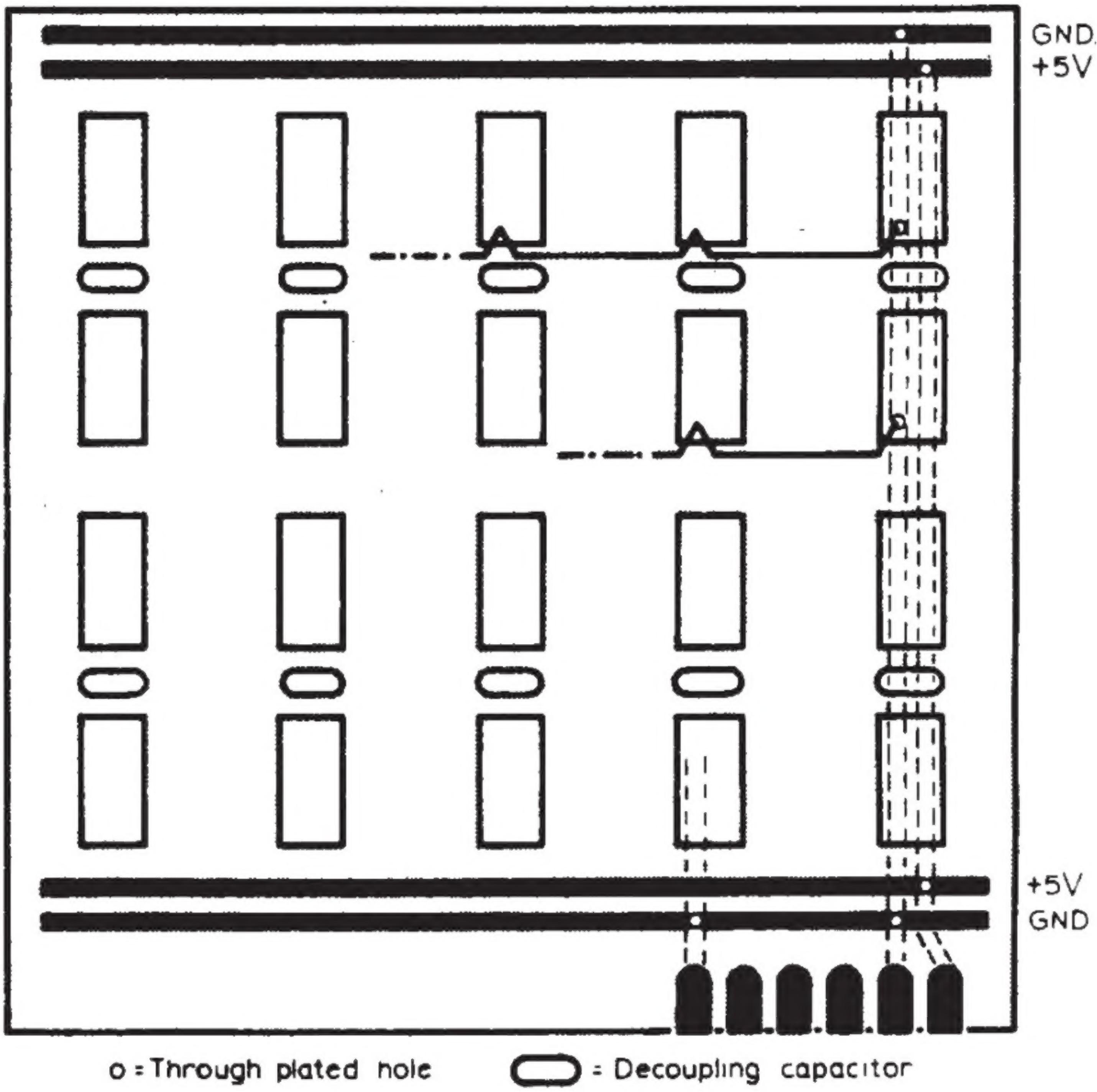


Fig. 7. Recommended layout.

Displacement current

— and how to get rid of it

by **I. Catt** and **M. F. Davidson** (CAM Consultants)
and **D. S. Walton** (Icthus Instruments Ltd)

To enable the continuity of electric current to be retained across a capacitor Maxwell proposed a "displacement current". By treating the capacitor as a special kind of transmission line this mathematical convenience is no longer required.

CONVENTIONAL electromagnetic theory proposes that when an electric current flows down a wire into a capacitor it spreads out across the plate, producing an electric charge which in turn leads to an electric field between the capacitor plates. The valuable concept of continuity of electric current is then retained by postulating (after Maxwell)[1] a "displacement current", which is a mathematical manipulation of the electric field *E* between the capacitor plates which has the dimensions of electric current and completes the flow of "electricity" (Fig. 1 (a) and (b)). This approach permits us to retain Kirchhoff's Laws and other valuable concepts, even though superficially it appears that at the capacitor there is a break in the otherwise continuous flow of electric current.

The flaw in this model is revealed when we notice that the electric current entered the capacitor at one point only on the capacitor plate. We must then explain how the electric charge flowing down the wire suddenly distributes itself uniformly across the whole capacitor plate. We know that this cannot happen since charge cannot flow out across the plate at a velocity in excess of the velocity of light. This paradoxical situation is brought about by a fundamental flaw in the basic model. Work on high speed logic design[2] has shown that the model of a lumped capacitance is faulty, and "displacement current" is an artefact of this faulty model.

The true model is quite different. Electric current enters the capacitor through a wire and then spreads out across the plate of the capacitor in the same way as ripples flow out from a stone dropped into a pond. If we consider only one pie-shaped wedge of the capacitor, as in Fig 1 (c), we can recognise it as a parallel plate transmission line whose only unusual feature is that the line width is increasing (and hence the impedance is decreasing). The

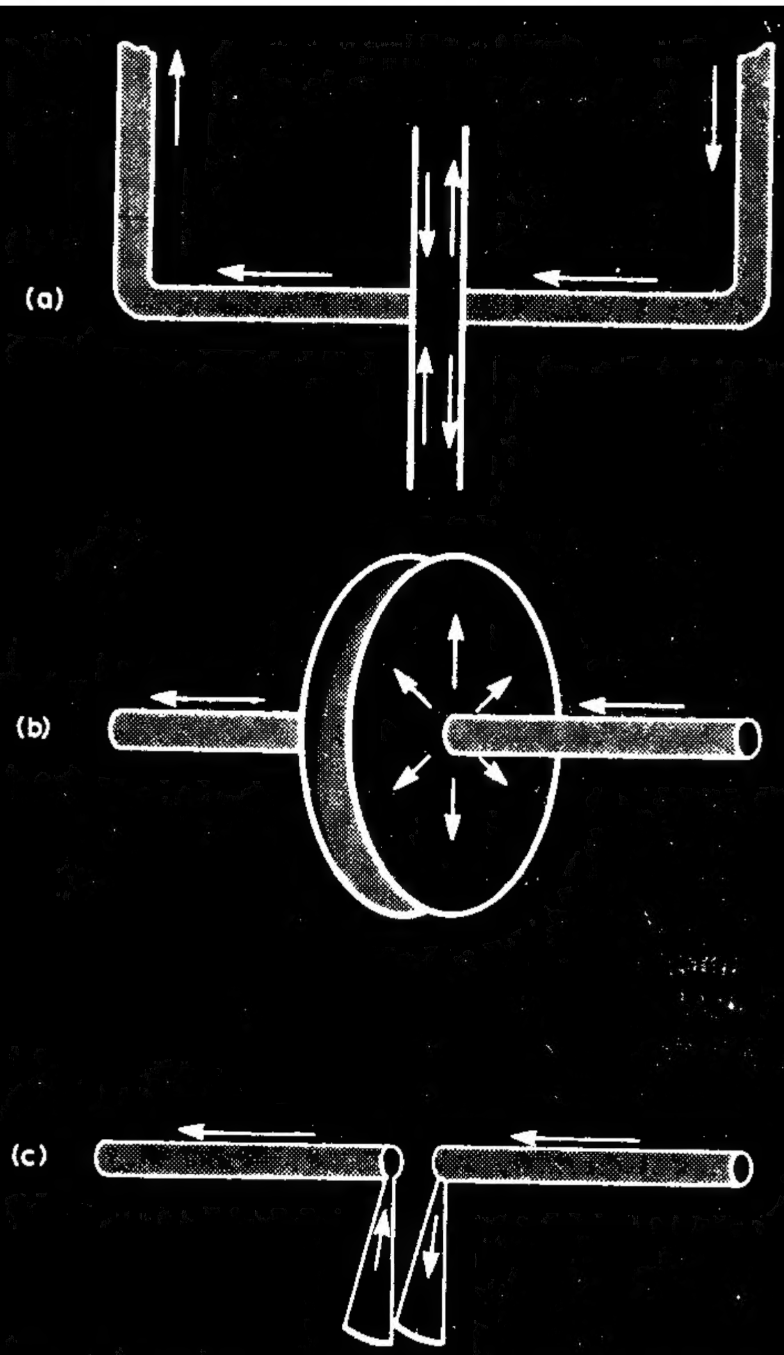


Fig. 1 *Process of current flowing into a capacitor and spreading out across a plate is shown in (a) and (b). The structure in (b) can be considered as being made up of a number of pie-shaped wedges as in (c), each of which is a transmission line.*

capacitor is made up of a number of these pie-shaped transmission lines in parallel, so the proper model for a capacitor is a transmission line.

Equivalent series resistance for a capacitor is the initial characteristic impedance of this transmission line at a radius equal to the radius of the input wires. Series inductance does not exist. *Pace* the many documented values for series inductance in a capacitor, this confirms experience that when the so-called series inductance of a capacitor is measured it turns out to be no more than the series inductance of the wires connected to the capacitor. No mechanism has ever been proposed for an internal series inductance in a capacitor.

Since any capacitor has now become a transmission line, it is no more necessary to postulate "displacement current" in a capacitor than it is necessary to do so for a transmission line. The excision of "displacement current" from Electromagnetic Theory has been based on arguments which are independent of the classic dispute over whether the electric current causes the electromagnetic field or vice versa.

Appendix

Comparison of the transmission line model with the lumped model of a capacitor in an RC circuit.

Taking the above discussion further, consider a transmission line as shown in Fig. 2, assumed to be terminated with a resistance R_T (not shown). The reflection coefficient is $\rho = (R_T - Z_o)/(R_T + Z_o)$ where Z_o is the characteristic impedance of the line. If the line is open-circuit at the right-hand end, as shown (and therefore R_T is infinite), the $\rho = +1$. We will assume that $R \gg Z_o$.

When switch S is closed (at time $t = 0$) a step of voltage $V.Z_0/(R+Z_o)$ is propagated down the line. This reflects from the open circuit at the right hand end to give a total voltage $2V.Z_o/(R+Z_o)$. Reflection from the left end makes a further contribution of $[V.Z_0/(R+Z_0)] \times [(R-Z_0)/(R+Z_0)]$ and so on. In general after n two-way passes the voltage after n passes is V_n and,

$$V_{n+1} = V_n + 2 . \frac{VZ_0}{R+Z_0}\left[\frac{R-Z_0}{R+Z_0}\right]^n \tag{1}$$

In order to avoid a rather difficult integration it is possible to sum this series to n terms using the formula,

$$= \frac{a(1-v^n)}{1-v} \tag{2}$$

where a is the first term of a geometrical progression and v the ratio between terms. (This formula is easily verified by induction.) Substituting in (2) the parameters from (1),

$$\text{i.e. } a = \frac{2VZ_0}{R+Z_0} \tag{3}$$

$$v = \frac{R-Z_0}{R+Z_0} \tag{4}$$

We obtain,

$$V_n = \frac{\frac{2VZ_0}{R+Z_0}\left[1-\left[\frac{R-Z_0}{R+Z_0}\right]^n\right]}{1-\frac{R-Z_0}{R+Z_0}} \tag{5}$$

$$= V\left[1-\left[\frac{R-Z_0}{R+Z_0}\right]^n\right] \tag{6}$$

This is the correct description of what is happening as a capacitor charges. We can now go on to show that it is approximated by an exponential. We have

$$V_n = V\left[1-\left[\frac{R-Z_0}{R+Z_0}\right]^n\right] \tag{7}$$

Consider the term,

$$T = \left(\frac{R-Z_0}{R+Z_0}\right)^n$$

$$= \left(\frac{1-Z_0/R}{1+Z_0/R}\right)^n$$

If $Z_0/R \ll 1$ this term is asymptotically equal to

$$\left(1-\frac{2Z_0}{R}\right)^n$$

Now define $k = 2Z_0 n/R$. Substitution gives:

$$T = \left[1-\frac{k}{n}\right]^n$$

By definition, as $n \to \infty$ we have,

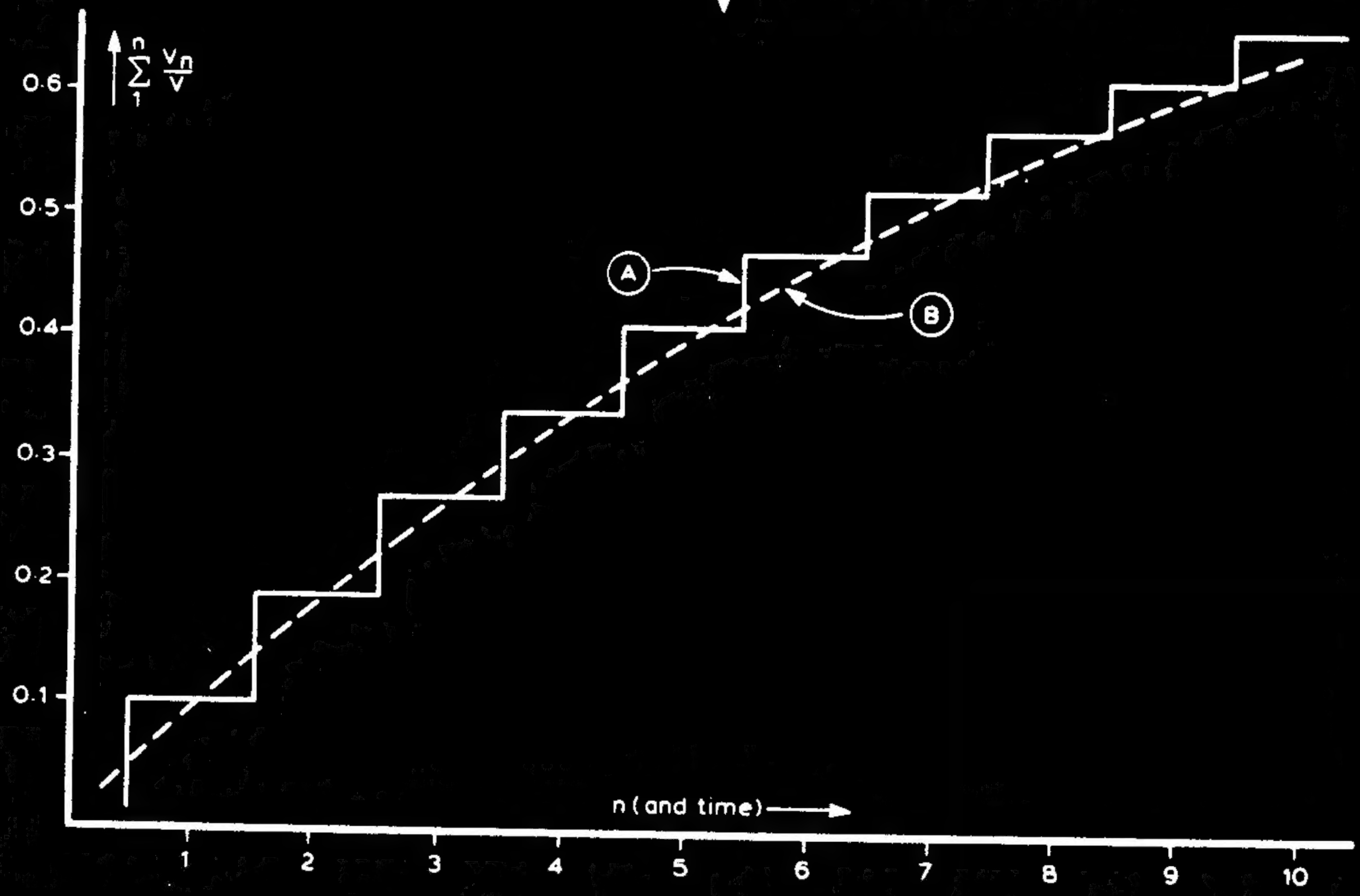
$$T = e^{-k} = e^{\frac{-2Z_0 n}{R}}$$

And therefore:

$$V_n = V\left[1-e^{\frac{-2Z_0 n}{R}}\right]$$

◄**Fig. 2** *An open-ended transmission line.*

Fig. 3 *Comparison of the transmission line model $1-(1-2Z_0/R)^n$ in the curve A with the lumped model $1-e^{-2Z_0 n/R}$ in curve B, for $2Z_0/R = 0.1$.*
▼



Now, after time t, $n = V_c t/2l$, where V_c = velocity of propagation.

Therefore

$$V(t) = V\left[1 - e^{\frac{-V_c t}{l} \cdot \frac{Z_0}{R}}\right]$$

For any transmission line it can be shown that:

$$Z_0 = f\sqrt{\frac{\mu}{\epsilon}}$$

$$V_c = \frac{1}{\sqrt{\mu\epsilon}}$$

$$C_1 = \epsilon/f$$

where C_1 = capacitance per unit length, and f is the same geometrical factor in each case. The "total capacitance" of length l of line $= l.C_l = C$.

$$\text{Hence}\ \frac{V_c Z_0}{lR} = \frac{1}{RC}$$

and therefore

$$V(t) = V(1 - e^{-t/RC})$$

which is the standard result. This model does not require use of the concept of charge. A graphical comparison of the results is shown in Fig. 3.

References

1. "History of displacement current", I. Catt, M. F. Davidson, D. S. Walton. *Physics Education*, to be published early 1979.

2. "Crosstalk (noise) in digital computers", I. Catt. *IEEE Trans.* EC-16, Dec. 1967, pp. 743-763.

DISPLACEMENT CURRENT

The explanation given by Messrs Catt, Davidson and Walton (December 1978, p.51) of the flow of current 'through' a capacitor without resorting to Maxwell's concept of displacement current is attractive to me, because notwithstanding my immense respect for Maxwell I have always felt that displacement current was a kind of subterfuge to get over a logical difficulty*. But before wholeheartedly accepting this alternative I would like to be given certain reassurances.

At the foot of column 1 the authors point out that the parallel elements of the disk capacitor depicted can be regarded as transmission lines whose characteristic impedance (Z_0) is continuously decreasing towards the far end. So there would be gradual reflection all the way. But in the mathematical proof Z_0 is treated as constant and there is reflection only at the far end. This made me feel I was being conned.

According to Ampère's Law, the connecting leads carrying the charging current must be everywhere encircled by a magnetomotive force numerically equal to the current. In the authors' Fig. 1 the leads are horizontal and the plates are in vertical planes, parallel to one another and also to the n.n.f. around the leads. But what about the m.m.f. in the space between the plates, due to what we have become accustomed to calling displacement current? This current, being a continuation across the capacitor gap of the external circuit current, one naturally sees its m.m.f. also as in a vertical plane. Can the authors show clearly how this follows from the geometry of their transmission line currents, which flow everywhere at right angles to the current in the leads? This aspect is of some importance, since the propagation of radio waves depends on it. Can the authors

convincingly get rid of displacement currents in space?

M. G. Scroggie,
Bexhill,
Sussex.

•But I never had, or heard of, a difficulty created by imagining current having to flow across the capacitor plates faster than light. Where did the authors get that idea? And why wouldn't it apply also to the current in the leads?

The authors reply:

The article discusses a circular capacitor. The appendix discusses a rectangular capacitor in order to minimize mathematical complexity. The appendix proves that if a voltage source is switched across a resistor and a rectangular capacitor in series, a waveform results which approximates to an exponential. As Mr Scroggie points out, it does not prove the same for a non-rectangular capacitor.

If you ask us to resolve paradoxes in classical theory, you are asking us to say that we are saying nothing that is fundamentally new; you are asking us not to publish anything. Do you believe that "new" information is only acceptable if it indicates no flaws in the conventional wisdom, i.e. if it is not really new?

As to the m.m.f. in the space between the plates, this have never been measured. If it had been measured it would have been found to be non-uniform, and the revered B. I. Bleaney and B. Bleaney ("Electricity and Magnetism," Clarendon, 1957, p.238) and others would not have written ". . . . the field in between the plates is uniform . . .", which of course it is not; a TEM waveform advancing between the plates of a capacitor (≡ transmission line) creates a field behind itself but not ahead of itself.

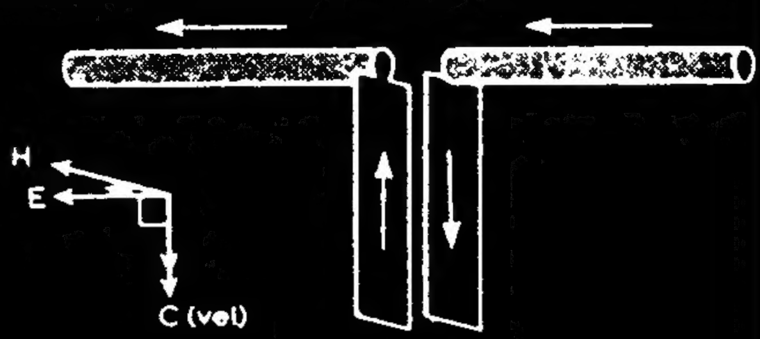
The last paragraph of Mr Scroggie's letter is crucial. If the capacitor were rectangular and oriented much as shown in our Fig. 1(c) then no m.m.f. in the vertical plane would result from current in the capacitor plates. Vertical m.m.f. would mean that the waveform was not TEM, but we know that it is TEM and travelling vertically downwards between the capacitor plates. That is, *E* and *H* fields are at right angles to the (downwards) direction of propágation, and therefore are horizontal. This is no more paradoxical than trying to apply Ampère's Law to a TEM step travelling along any transmission line.

Ampère did not know that a TEM wave ($E \times H$) travels forward between two wires at the speed of light. He did not know that a capacitor is a transmission line; he did not know about transmission lines.

These matters will be discussed further in a forthcoming article in *Wireless World*. A paper "The Heaviside Signal" will further clarify the situation (see "Electromagnetic Theory Vol 1," published by C.A.M. Publishing, 17 King Harry Lane, St Albans).

I. Catt, M. F. Davidson, D. S. Walton

Further letters on this subject will be published later. Ed.



(See Electronics and Wireless World, Jan.88, page 54.)

MAXWELL, EINSTEIN AND THE AETHER

I regret that lack of space prevented me from publishing the above forthcoming article in this book.

The summary of the article is as follows;

> "To sum up. Einstein says that* relativity, which he believes to have been based on the disappearance of a space with physical properties, is based on Maxwell's Equations, which are now [see pp 187,138] found to contain <u>only</u> information about the physical attributes of that disappearing space.
>
> "By analogy, it would be possible to proclaim a new theory of mechanics which lacked the concept of mass, but which contained both velocity v and momentum mv within it, and which preferably included lots of fancy maths involving momentum and velocity. .... The necessary parameter m, like the rabbit in the hat, could go about its business... firmly hidden in ... a fog of mathematics..."

* See footnote, page 193

C.A.M.

The Heaviside signal

An alternative view of the transverse electromagnetic wave

by I. Catt, CAM Consultants

This article proposes a different picture of electromagnetic propagation from the familiar "rolling wave" idea in which electric and magnetic fields topple over and forward, continually changing into each other as they go. The author postulates "an unchanging slab of E x H energy current" travelling forward at the speed of light, and names it "the Heaviside signal" after a concept expressed in the writings of Oliver Heaviside. This process does not rely on a causal relationship between the electric and magnetic fields, which are seen as co-existent.

MAXWELL faced up to the paradox that whereas electric circuits, in order to function properly by allowing the passage of electric current, were thought to require a complete closed circuit of conductors, electric current still seemed to flow for a time when a capacitor (which of course is an open circuit) was placed in series with the closed loop of conductors. He "cut the Gordian knot" (according to Heaviside)[1] by postulating that a new kind of current, which he called "displacement current", leapt across the plates within the capacitor. This electric current, which was uniformly distributed in the space between the capacitor plates, could even flow through a vacuum.

Maxwell followed up this daring idea by suggesting that electromagnetic waves might exist in space. Scepticism about his postulated "displacement current" was silenced in 1887 when Hertz discovered the predicted waves in space. The classic pre-Popperian requirement of a good scientific theory seemed to have been met – the prediction of further results which are later confirmed by experiment.

There are two versions of the transverse electromagnetic wave, the "rolling wave," and what we shall call here the "Heaviside signal." We shall discuss only the wide variety of views among those who believe (with the relativists) that there is no instantaneous action at a distance.

The rolling wave

The lack of action at a distance creates a fundamental difficulty for the wave in space if it is to be launched by a force in the direction of propagation. The key to the ability of a force to project a wave is that there is a pressure difference between two points along the line of propagation. However, knowledge of a difference in pressure between two points A and B which are separated by distance implies instantaneous knowledge at B of the pressure at A; that is, instantaneous action at a distance, which has been outlawed.

This dilemma seems to be overcome if it is postulated that the force which projects the wave is a lateral, shear, force. It seems a shear force can act at a point, and so not contradict relativity whereas a longitudinal force cannot.

The above kind of reasoning, combined with the postulation of displacement current, which seemed to flow at right angles to the direction of propagation, joined forces to create the notion of the rolling wave. The rolling wave contains alternating concentrations of magnetic energy $\frac{1}{2}\mu H^2$ and electric energy $\frac{1}{2}\epsilon E^2$ in the direction of propagation. It is useful to think of a road with alternate red trucks and white motor cars. The magnetic energy or flux (by Faraday's law of induction) generates electric energy and displacement current ahead of itself, which in turn (by the Biot-Savart Law) generates magnetic flux, or energy, ahead of *itself*. Each type of energy, or flux, topples over and forward, changing as it topples into the other kind of energy. It is as though in the road containing the alternate red trucks and white cars, first the red trucks reappear as white cars a little further ahead while at the same time the white cars turn into red trucks a little further ahead; then the trucks and cars change back again, moving forward a little with each metamorphosis. The analogy with the pendulum has been proposed. One can think of a long line of pendulums, alternate ones having potential energy and kinetic energy, and communicating their energy forward step by step with a change of type of energy at each one.

The Heaviside signal

Opposed to the rolling wave is what we have called the Heaviside signal. The most highly developed form of this view is that at any point in space, an electromagnetic signal always contains one kind of energy only, which is equal to the product of E and H at that point, where

$$\frac{E}{H}=\sqrt{\frac{\mu}{\epsilon}}. \qquad \text{Energy density} = \frac{E.H}{c}$$

Further, the Heaviside signal always travels forward unchanged at the speed of light, $c = 1/\sqrt{\mu\epsilon}$, and never any slower. E, H and c are always mutually perpendicular.

The two men most likely to understand the "Heaviside signal" point of view and to oppose the "rolling wave" were Oliver Heaviside himself, in honour of whom it has been given its name, and Poynting, the man whose name is attached to the vector $E \times H$. However, their writings show that neither man arrived at a full understanding of the Heaviside signal described in the previous paragraph.

Heaviside vacillated between the two views, the rolling wave and the Heaviside signal. He always applauded the idea of displacement current, which appears to put him on the side of the rolling wave. Further, on page 6, art. 453 of volume 3 of his "Electromagnetic Theory", when he says that the curl of E, not E itself, is the real source of the waves, he is again arguing for the rolling wave. Curliness is obviously a bid for shear, vorticular forces, a concept intrinsic to the rolling wave. However, elsewhere he seems to stand firmly for the Heaviside signal. For instance *(ibid*, art. 451, page 4), he says, "It carries all its properties with it unchanged," which is a clear statement of the Heaviside signal. In art. 452, the mention of a "slab" of signal is strongly on the side of the Heaviside signal. Heaviside mentions the slab elsewhere in his writings. One does not conceive of slabs rolling, or generating shear forces or stresses. Almost by definition, a slab, like a slab of heavy granite, moves forward unchanged at constant velocity.

Professor Poynting, who first suggested that energy was distributed in space with a density $E\times H$, also had a partial vision of the Heaviside signal. He definitely did not know that E is always perpendicular to H, and that the $\times$ in $E\times H$ means simply multiplication. (He had a term $\sin\theta$ for the angle between them.) Poynting was writing before the general agreement that light is electromagnetic, and so did not know that this Poynting energy $E\times H$ always moved forward (in the third dimension) at a constant speed, $1/\sqrt{\mu\epsilon}$, the velocity of light in the medium.

Poynting had a very good grasp of the direction of energy flow and its magnitude, but did not seem to understand the importance of reflections at a change of medium, which leads one to think of one energy current $E\times H$ flowing backwards along its previous path, passing through the next portion of forward travelling energy current. This superposition of forward and backward energy currents (implicit in the phrases "phase velocity" and "group velocity") has prevented a clear understanding of the electromagnetic wave.

For fifty years, technology did not give us the power to drive the medium with an electromagnetic signal. With the low power at our disposal, all we could do was resonate the medium with periodic (sinusoidal) excitation in the same way as we move a child on a swing. In a resonant medium, energy is necessarily flowing in both directions; most of the forward energy returns to aid the source on the next cycle.

Our inability to drive a medium except periodically insinuated itself into our group psyche, until we came to assert that nature *was* periodic (and even that it was sinusoidal). Implicit in this view were the wrong beliefs that

(1) electromagnetic energy is necessarily contrapuntal,
(2) $E/H=\sqrt{(\mu/\epsilon)}$ is not always true, (e.g. when two waves are passing through each other so that H cancels but E does not, so that $E/H=\infty$), and
(3) signals can travel slower than the speed of light $1/\sqrt{\mu\epsilon}$).

The absurdity of this third idea is easy to demonstrate if we consider a two directional highway. If all cars move at 60 m.p.h. but some (A per hour) move eastwards and some (B per hour) move westwards, no one would argue that the total passage of cars eastwards per hour past a reference point, that is, $(A-B)$, would help us to determine the velocity of cars by the formula
Flow of cars $=(A-B)$ per hour
Distance between cars $=L$
Therefore velocity of cars $=(A-B)L$ m.p.h.
However, this seems to be done, at least subconsciously, with phase velocity and group velocity. The very terms imply some such calculation.

Some ten years ago the successful manufacture of high speed (1ns) logic elements capable of driving a 100 ohm load made it possible, for the first time for fifty years, to drive a medium rather than gently resonate it, as a matter of normal routine. Those driving a high speed logic step could clearly see it travelling at the speed of light for the dielectric (never any slower) and remaining unchanged on its journey. For the first time for seventy years, high speed digital engineers were privileged to see the Heaviside signal, an unchanging slab of $E\times H$ energy current guided between two conductors from one logic gate to the next. Reflections were prevented by proper termination at the destination, so that notions of phase velocity and group velocity

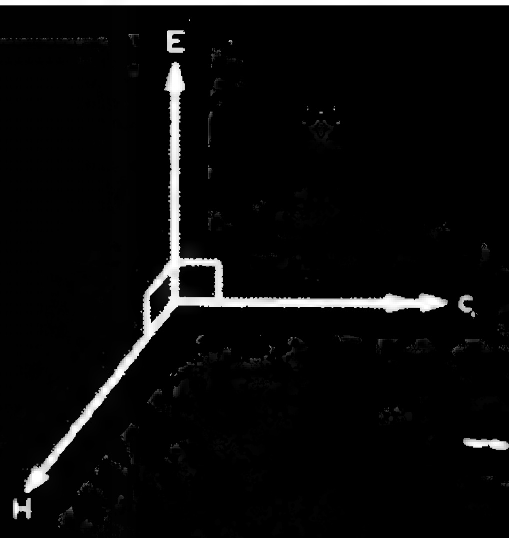
evaporated. We saw a slab of energy launched from one point, travelling unaltered, to be absorbed by the terminating resistor at the destination.

At this point we just *had* to unburden ourselves at the theoretical level of implicit contrapuntal notions. A beautiful vision resulted, now called the Heaviside signal, of a lateral strain $E \times H$ (where $E/H = \sqrt{\mu/\epsilon}$ which by definition travelled forward at velocity $1/\sqrt{\mu\epsilon}$. As it travelled forward it filled (or probed) the space ahead of it in the same way as the ripples on the surface of a pond will fill the space (surface) as they come to it. Logic designers maintained a near constant aspect ratio in the space ahead, because whenever this slab came to a change in aspect ratio (= change of characteristic impedance, better termed characteristic resistance) some of the energy current would double back on its tracks according to the well-known laws of reflection. However, this did not lead back to the old "phase velocity" and "group velocity" notions; rather, the slab of energy current split into two slabs, one to continue forward and the other to return, both slabs continuing to probe, or fill, the space presented to them on their journeys.

The Heaviside signal offers us a dramatic simplification of our view of the fundamentals of electromagnetic theory.

Definitions

First define energy current (= TEM wave = Poynting vector) as our primitive, where energy current is as follows:



Now $\sqrt{\mu/\epsilon}$ and $1/\sqrt{\mu\epsilon}$ can be independently defined. Let us define

(a) $\sqrt{\frac{\mu}{\epsilon}} = \frac{E}{H}$

which defines a constant of proportionality for the medium.

(b) $\frac{1}{\sqrt{\mu\epsilon}}$ = velocity of propagation c,

again a constant for the medium.

(c) Define $D = \epsilon E$, $B = \mu H$

Derivations

$$\frac{E}{H} = \sqrt{\frac{\mu}{\epsilon}}, \quad B = \mu H$$

$$\therefore \frac{E\mu}{B} = \sqrt{\frac{\mu}{\epsilon}} \qquad (1)$$

$$\therefore \frac{E}{B} = \frac{1}{\sqrt{\mu\epsilon}} = c \qquad (2)$$

$$\therefore E = Bc \qquad (3)$$

By definition*,

$$c\frac{\partial E}{\partial x} = \frac{\partial E}{\partial t} = -c\frac{\partial B}{\partial t} \qquad (4)$$

*See Appendix 1

$$\frac{\partial E}{\partial x} = -\frac{\partial B}{\partial t} \qquad (5)$$

This is equation (12.5.1) in Carter (G. W. Carter, The Electromagnetic Field in its Engineering Aspects, Longmans, 1954, page 268), when he believes he is *deriving* the TEM wave, which is supposed to *result* from a causality relationship between E and B (Faraday's law of electromagnetic induction). Carter is clearly developing the rolling wave.

We see then that the equation $\partial E/\partial x = -\partial B/\partial t$ is a simple derivation from the definition of the Heaviside signal and is not based on $\partial B/\partial t$ *causing* E, as Faraday thought he had discovered.

We have shown that the passage of a TEM wave and all the mathematics that has mushroomed around it does not rely on a causality relationship (or interchange) between the electric and magnetic field. Rather, they are coexistent, co-substantial, co-eternal. The medium can only be strained in the two lateral dimensions (E and H) in fixed proportion. [In a similar way, *pressure* in a liquid in direction x does not *cause* pressure in the y (and z) direction; they co-exist.]

Faraday's great discovery in the 1830s was *not* electromagnetic induction; not a causality relationship. His great achievement was to discover that *change* was important. This started us on the road to discovering the now postulated primitive, the Heaviside signal, which can only move; it cannot stand still. Heaviside put together the main features of the new concept, but it took another century to put flesh on to the bare bones.

Reference

1. Oliver Heaviside, Electromagnetic Theory, 1893, London, page 28 section 30.

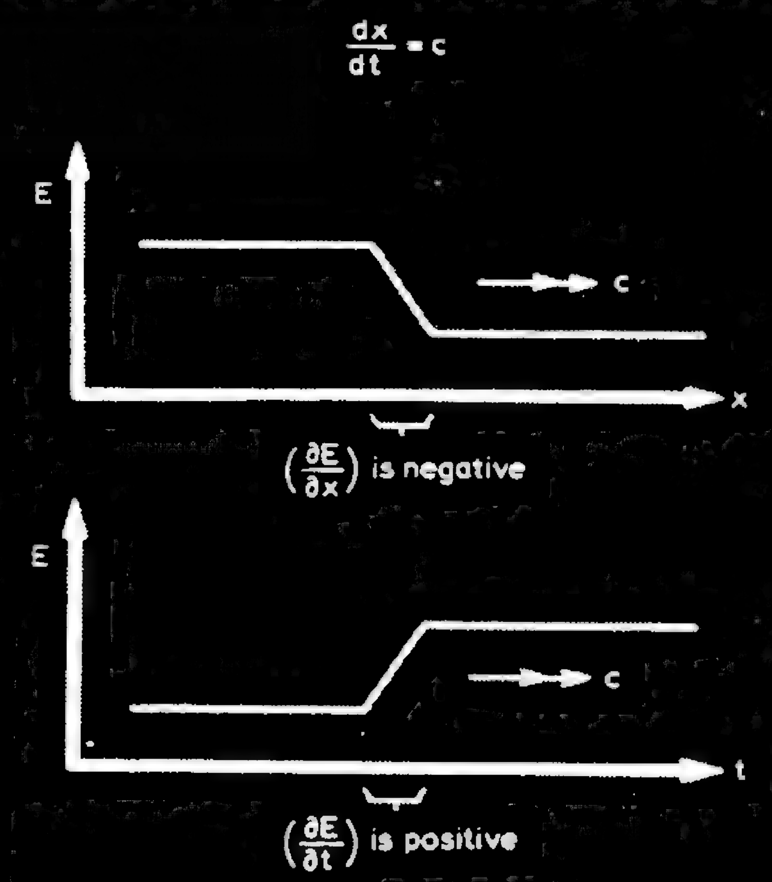
Appendix 1

By convention, if a voltage step is travelling from left to right (i.e. in a positive direction) it has a positive velocity; dx/dt is positive

$\frac{\partial E}{\partial t}$ is positive but $\frac{\partial E}{\partial x}$ is negative. This

(reversal) problem is well known by anyone who has drawn out an oscilloscope trace on to paper with voltage and distance axes. This explains the minus sign in equation (4) in the article. When we travel, we *gain* distance while we *lose* time. However, we regard our velocity dx/dt as positive.

It is strange that this ambiguity in sign convention had led to a negative sign in electromagnetic theory. This in turn intro-



duced the idea of a "Lenz's law" reluctance, or back e.m.f., in which lies nested the idea of causality.

$$i \rightarrow \int H dl \qquad \text{and} \frac{dB}{dt} \rightarrow v$$

In fact, electric and magnetic fields have a positive relationship, and co-exist rather than cause each other.

Numerically,

$$c\left|\frac{\partial E}{\partial x}\right| = \left|\frac{\partial E}{\partial t}\right|$$

Therefore, since by convention $\partial E/\partial t$ is positive, $\partial E/\partial x$ is negative and c is positive, we must conclude that

$$c\frac{\partial E}{\partial x} = -\frac{\partial E}{\partial t}$$

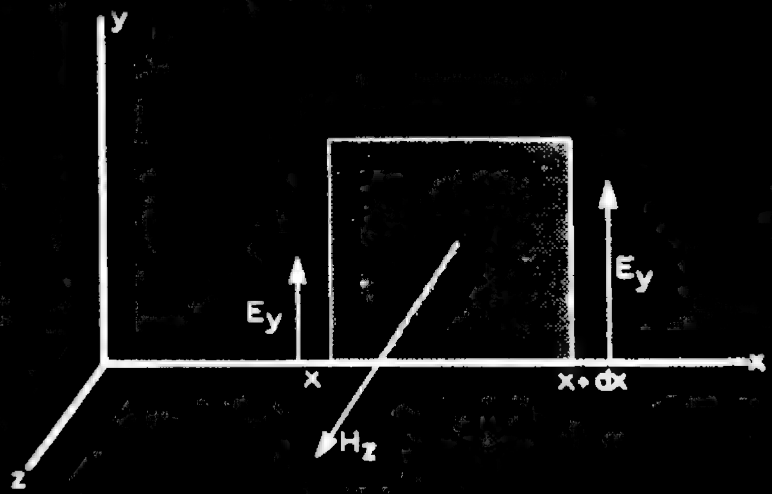
Appendix 2: the rolling wave explained

In this article, two mutually contradictory versions of the transverse electromagnetic wave have been described and compared. These were the rolling wave and the Heaviside signal. This appendix contains the first half of a very clear description of the rolling wave taken from "Fundamentals of Electricity and Magnetism" by Arthur F. Kip, Professor of Physics, University of California, Berkeley, published by McGraw-Hill, 1962, page 320. Only enough of that description is reproduced to make his approach clear.

"... Our demonstration involves the use of the first two Maxwell equations to show that such a postulated time and space variation of *E* gives rise to a similar time and space variation of *H* (but at right angles to *E*) and that this *H* variation acts back to cause the postulated variation in *E*. Thus, once such a wave is initiated, it is self-propagating.

"The figure below is used to show the application [of Faraday's law of induction] to the plane *E* wave, postulated to be moving along the *x* direction. A convenient closed path is drawn in the *xy* plane, around which we shall take the line integral of *E*. This is equated through [Faraday's law] to the rate of change of flux *H* through the plane bounded by the path of the line integral. Only the vertical parts of the line integral contribute since *E* is in the *y* direction, so that $E.\partial x = 0$. If we go around in a counterclockwise direction, the line integral around the path chosen becomes

$$\oint E \cdot dl = (E_y)_{x+dx}dy - (E_y)_x dy$$
$$= [(E_y)_{x+dx} - (E_y)_x]dy$$



where we are to take the values of E_y at $x+dx$ and x, respectively. The difference between these two values of E_y at the two positions is $(\partial E_y/\partial x)dx$, so we can write the line integral of Faraday's law of induction as

$$\frac{\partial E_y}{\partial x} dxdy = -\mu_0 \frac{\partial H_z}{\partial t} dxdy$$

Since this relationship is true for any area dxdy, we may write

$$\frac{\partial E_y}{\partial x} = -\mu_0 \frac{\partial H_z}{\partial t}$$

(This ends the extract from Kip. To get to the Carter equation we have to replace μH by *B*, of course.)

DISPLACEMENT CURRENT

The two articles on displacement current which have recently appeared in your magazine (December 1978, March 1979), contain the sensible suggestion that one should regard currents and charge distributions as the consequences of electromagnetic waves rather than as the sources of these waves. Apart from this, the articles are wrong in almost every detail and it is vital that this should be clearly demonstrated before undue damage is done.

The basic demolition process is simple. In Maxwell's equations for a dielectric medium we have,

$$\operatorname{div} \mathbf{D} = 0, \quad \operatorname{div} \mathbf{B} = 0,$$

$$\operatorname{curl} \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}, \quad \operatorname{curl} \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},$$

Writing $\mathbf{D} = \epsilon \mathbf{E}$ and $\mathbf{B} = \mu \mathbf{H}$ for a linear, homogeneous, isotropic medium, these equations give the wave equation for **E** (or **H**),

$$\nabla^2 \mathbf{E} = \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2}$$

which means that electromagnetic waves travelling with a speed of $1/\sqrt{\mu\epsilon}$ can exist in the dielectric. The wave equation occurs due to the presence of the term $\partial \mathbf{D}/\partial t$, which Maxwell introduced and called "displacement current". Without this term the wave equation would not appear and electromagnetic waves would not exist. There is a fair amount of evidence that electromagnetic waves *do* exist, and I doubt if Catt, Davidson, and Walton would deny this. I would like to believe that they are only objecting to the name "displacement current", but if that were the case there would hardly be any point in making such a vicious attack, and after re-reading these remarkable essays a number of times I have a feeling that C, D, and W really believe that electromagnetic waves can exist without $\partial \mathbf{D}/\partial t$ occurring in the equations. A consultation with any competent mathematician should convince them that this is not so.

The above argument may not be very convincing to the non-mathematical reader and perhaps C, D, and W won't like it very much, because one gets the strong impression that these gentlemen have probably used Maxwell's equations in only the most trivial of problems. It is therefore necessary to criticise the articles in some detail. Take, for example, the simple reflection treatment given in the appendix to the first article. This applies to a uniform transmission line but not, as stated in the appendix, to a non-uniform line. For a uniform line the wave equation is

$$\frac{\partial^2 v}{\partial x^2} = LC \frac{\partial^2 v}{\partial t^2} = \mu\epsilon \frac{\partial^2 v}{\partial t^2},$$

where L and C are the inductance and capacitance per unit length. The error probably arises due to the following plausible but erroneous argument: "In the circular capacitor L and C vary with r,

$$L = \mu \frac{d}{2\pi r}, \quad C = \epsilon \frac{2\pi r}{d}$$

Hence the product LC is still constant and equal to $\mu\epsilon$. So the wave equation for the circular capacitor will be

$$\frac{\partial^2 v}{\partial r^2} = \mu\epsilon \frac{\partial^2 v}{\partial t^2}"$$

If the wave equation is properly derived from the basic equations it will be found to be

$$\frac{\partial^2 v}{\partial r^2} + \frac{1}{r}\frac{\partial v}{\partial r} = \mu\epsilon \frac{\partial^2 v}{\partial t^2}$$

The reason why the reflection process described for the uniform line does not apply in this case is that there is a continuous reflection from the wave front due to the con-

tinuous variation of Z_0. Another serious error is that the authors regard the "radius of the input wires" to be the "input end" of the circular transmission line. If they had taken the trouble to consider the Poynting vector field, they would have discovered that the energy enters the capacitor dielectric at the outside radius, and that this outside radius is the input to the capacitor. When they take a sector of this capacitor (Fig. 1(c) of the first article) they *do* have a line supplied at the inner radius. Hence it is incorrect to regard the complete capacitor as a large number of such sectors ("pie-shaped"!) in parallel.

In the second article, and also in their reply to Mr P. I. Day's sensible letter, the authors ask "where, then, is the displacement current in the transmission line?". The answer, of course, is that in general it flows in all parts of the dielectric, but by choosing a "step" wave (a physical impossibility) they have pushed all of the displacement current into an infinitely thin sheet in the wavefront and have lost sight of it. But we haven't. A step is a very useful concept as the limiting case of, say, an exponential rise, but if the limiting process is improperly understood and causes one to lose things, it is advisable not to use it. And do I detect a rather nervous reaction to Mr Day's use of the frequency domain? Did they for one awful moment think that they saw the ghost of Maxwell's displacement current? They need not worry, it is not dead yet and *they* are certainly not capable of killing it.

These three gentlemen see fit to criticise Maxwell for lack of insight, and assert that Maxwell did not realise that displacement current was not uniformly distributed within a capacitor. In other words, that he was not capable of getting the correct solution to his own equations! And finally they praise Heaviside for "missing it only by a whisker". In fact Heaviside was never in any such danger, but I am afraid that Catt, Davidson, and Walton have dropped right in it!

May I suggest that your readers will be well advised to approach the "further reading" with great caution.

B. Lago
Doxey
Stafford

The authors reply:
Dr Lago's letter raises some interesting points which probably deserve fuller treatment than we are able to give here. We are interested that he should feel that "undue damage" can be done to Maxwell's theory through this series of articles. It would seem that he sees himself in the role of priest defending the faithful from the dangers of heretical doctrine. If this is indeed necessary then it says little for the understanding of electromagnetic theory by the faithful. Surely engineers and scientists are competent to draw their own conclusions from a public debate without such protection.

Dr Lago states "Without this term (displacement current) the wave equation would not appear and electromagnetic waves would not exist". Would that life were so simple! In fact this statement is a non-sequitur. All that he is able to state from his position is something like, "In Maxwell's theory displacement current is essential to the existence of a wave equation and hence of electromagnetic waves; therefore, if displacement current is removed, electromagnetic waves *as understood by Maxwell* would not exist". To illustrate, before Lavoisier it was thought that the process of combustion involved, or rather depended upon, the removal of a substance, 'phlogiston', from the burning material. Someone who believed the phlogiston theory would no doubt have asserted that "without phlogiston it is impossible for things to burn". But he would have been quite wrong because the argument is premised on a faulty theory. In the same way we regard the Maxwellian framework as faulty. We have no doubt that electromagnetic radiation exists and there is nothing in our articles to suggest otherwise. What we chiefly object to is the spurious causality and physical meaning given to the term $\epsilon(\partial E/\partial t)$ which is a barrier to the deeper understanding of electromagnetic processes.

We would like to assure Dr Lago that our experience in electromagnetic theory goes beyond "the most trivial problems" and one of us (DSW) lectured on electromagnetic theory in Trinity College Dublin.

Dr Lago is quite wrong to impute to us the facile misunderstanding of the pie-shaped transmission line. IC published a paper[1] in

which the theory of the pie-shaped line is discussed with reference to power plane decoupling on multi-layer printed circuit boards. In this paper it is made quite clear that there is continuous reflection caused by the changing impedance seen by the step as it travels outwards to greater radii. We did in fact reference this paper at the end of the December 1978 article. In this latter article we do not claim to be treating the case of a circular capacitor in the mathematical appendix. We in fact refer to Fig. 2 which represents a uniform end-fed transmission line. This case is treated since it demonstrates the key features without requiring unnecessarily complex mathematics.

Incidentally, Dr Lago says that a zero risetime step is a "physical impossibility". This interesting statement merits further analysis. One would like to know whether he is attacking the concept or its practical realisation, i.e. is he against the Platonic ideal of a step or is he saying, as might Aristotle, that such a concept is not useful because it is not practically realisable? If the former then we assume he is also opposed to the sine wave concept since infinite time is required for its perfect realisation; if the latter then what physical principle determines the shortest risetime obtainable in practice? In the latter case the principle must precede the concept, i.e. there must be no circularity.

Finally, Dr Lago agrees with us (and Heaviside) when he states that "one should regard currents and charge distributions as the consequences of electromagnetic waves rather than as the sources of these waves." In that case is $\epsilon(\partial E/\partial t)$ a *current* and therefore an effect or a *field* and therefore a cause, or is it both!

I. Catt, M. F. Davidson, D. S. Walton

Reference

1. "Crosstalk (noise) in digital computers", I. Catt, *IEEE Trans. EC-16*, Dec. 1967, pp. 743-763.

C.A.M.

No radio without displacement current

An aid to understanding Maxwell's equations for wave propagation

by **D. A. Bell**, M.A., B.Sc., Ph.D., F.Inst.P., F.I.E.E.

"Faraday's conception of electric and magnetic force and their interrelations, expressed in terms of his lines of force, were fundamental. In terms of them James Clerk Maxwell developed the equations that underlie all modern theories of electromagnetic phenomena."
Encyclopedia Britannica.

BECAUSE displacement current forms a vital link in Maxwell's equations for wave propagation in empty space, text books often give the impression that Maxwell invented displacement current as a kind of mathematical trick to make his equations work. This is not so. In his two-volume *Electricity and Magnetism*, displacement current appears first on p.65 in volume 1, in the part dealing with electrostatics, and the idea follows from Faraday's work on lines of force. It is easy enough to think of electric and magnetic fields as stresses in a tangible medium such as insulating material or iron, but what happens when the material medium is replaced by a vacuum, leaving the fields 'hanging in space'? We no longer believe in an all-pervading ether, yet experience has long shown that light from the stars travels freely through space which is practically empty and now radio waves travel back to the earth from a vehicle which is near Jupiter. So it seems that we must accept that electromagnetic fields can exist in empty space.

But has something been slipped through in the last sentence? How did electric and magnetic fields come to be replaced by electromagnetic fields? Of course it was Maxwell who transformed "electricity and magnetism" into "electromagnetism" by setting out four equations which link together to form a closed cycle of electric field – magnetic field – electric field . . . and so on, continuing for ever as radiation if no conductors get in the way. Looked at from the experimental viewpoint, the most basic factor is electric charge, which usually is associated with a number of electrons (negative charge) or of protons (positive charge). A charge in steady motion constitutes a current, which produces a steady magnetic field. With varying motion a varying current produces a varying magnetic field which acts on an electric charge like an electric field. This looks